

## **Automatic, Profile-Free Web Page Recommendation**

### **NOTICE OF COPYRIGHTS AND TRADE DRESS**

**[0001]** A portion of the disclosure of this patent document contains material, which is subject to copyright protection. This patent document may show and/or describe matter, which is or may become trade dress of the owner. The copyright and trade dress owner has no objection to the facsimile reproduction by any one of the patent disclosure, as it appears in the Patent and Trademark Office patent files or records, but otherwise reserves all copyright and trade dress rights whatsoever.

### **RELATED APPLICATION INFORMATION**

**[0002]** This application is related to and claims priority from US Application No. 60/226,341, filed August 18, 2000, entitled "Automated Internet Touring System Tailored To User-Specific Qualities," which is incorporated herein by reference.

## **BACKGROUND OF THE INVENTION**

### **Field Of The Invention**

**[0003]** The present invention relates to providing content of interest to a user. The present invention further relates to recommending web pages to a user of the World Wide Web based upon the currently viewed web page.

### **Description Of Related Art**

**[0004]** The World Wide Web is a collection of millions of linked web sites, electronic documents and files that are stored on computers throughout the world. The World Wide Web includes Web sites that literally relate to millions of different subjects, which may or may not be of interest to a person who is surfing the web.

**[0005]** A user typically employs a browser to access Web sites that are of interest to the user. The user can learn of Web sites of interest by either learning of the site through friends or through the media. Another way to learn of Web sites is to use a search engine to search the Web. The user typically types key words into a search engine Web page. The search engine then returns a list of one or more Web sites that relate to the keywords. This can be confusing for novice users who are unfamiliar with computers and the Web. Moreover, the use of search engines can also be frustrating for experienced users as the search engines may sometimes turn up sites that are unrelated to the keywords. Even worse, the search engines may sometimes inadvertently turn up sites that are of an objectionable nature to a particular user.

**[0006]** It would be desirable for a program or an online service to automatically assist a user with browsing to Web sites that are particularly tailored to the user's interests. This would allow novice user to quickly become accustomed to using the Web in a relatively easy manner. Such a program or service would also provide experienced users with a more fulfilling online experience.

**[0007]** Several attempts have been made at providing users with automated browsing assistance. In a system called "Ringo" developed at the MIT Media-Lab in the mid-1990s, personalized recommendations were made to a user based upon similarities between the interest profile of that user and the interest profiles of other users. Ringo was designed for making recommendations of music albums and artists, though it applied to Web browsing. In Ringo, the user profiles were developed by having the user rate content.

**[0008]** Other browsing aids, such as the eTour service of eTour, Inc., also depended on the development of user profiles. The quality of profile-based services depends on the extent and accuracy of each user's profile. Thus, in some services, a considerable number of users, providing considerable amount of ratings, are required before they become useful. Furthermore, profile-based services cannot easily account for changing tastes of the users. Finally, profile-based services face a considerable obstacle in that, before a user can see the benefits of the service, the user must register and provide a profile. Many users prefer to browse anonymously, and studies have shown that users have relatively short attention spans.

[0009] Prior art content location aids are typically server-based. For example, the eTour service requires the user to register with their server, and the user must visit the eTour site each time a user wishes to activate the service during a session. Other server-based aids have been provided in web sites which allow users to make purchases from an on-line catalog. For example, in some web sites, when a user identifies a particular item in the catalog of interest, then the server, when dynamically creating a web for the user, may identify other products in the catalogue which may be of interest to the user. Such server-based aids are limited, in that they only work with a single on-line catalogue, and require that the user remain in contact with the server. These server-based aids can be slow, both because of the demands placed upon the server, and the need to make repeated data transfers over the telecommunications infrastructure.

2025 RELEASE UNDER E.O. 14176

**SUMMARY OF THE INVENTION**

**[0010]** In accordance with the present invention, an electronic content recommendation service is provided which can act as an aid to a user in obtaining electronic content. The service is provided using software, apparatus and methods in accordance with the invention. The service may be operated without user profiles or user registration. However, the service nonetheless can provide highly useful recommendation for electronic content to browse.

**[0011]** In accordance with the invention, a user's browsing of electronic content is monitored. For each unit of electronic content output by the user's browser, one or more units of electronic content (e.g., web pages) are recommended to the user. The user may then load a recommended unit of electronic content.

**[0012]** Recommendation is based upon a system of categorization. A number of units of electronic content are identified as fitting into predefined categories of human interest. During a user's browsing, the unit of electronic content loaded in the user's browser is determined to be in at least one of the predefined categories. Recommendations of electronic content to browse are drawn from lists of units of electronic content which were previously placed into the category of the current unit of electronic content.

## DESCRIPTION OF THE DRAWINGS

[0013] Further objects of this invention, together with additional features contributing thereto and advantages accruing therefrom, will be apparent from the following description of an embodiment of the present invention which is shown in the accompanying drawings with like reference numerals indicating corresponding parts throughout and which is to be read in conjunction with the following drawings, wherein:

[0014] Figure 1 is a first block diagram of a network data distribution system in accordance with the invention.

[0015] Figure 2 is a second block diagram of the network data distribution system in accordance with the present invention.

[0016] Figure 3 is a representation of a display of a local device having a client window and a browser window in accordance with the invention.

[0017] Figure 4 is a flow chart of a method of recommending web pages to a user in accordance with the invention.

[0018] These and additional embodiments of the invention may now be better understood by turning to the following detailed description wherein an illustrated embodiment is described.

**DETAILED DESCRIPTION OF THE INVENTION**

**[0019]** Throughout this description, the embodiments and examples shown should be considered as exemplars, rather than limitations on the apparatus and methods of the present invention.

**The System and Apparatus of the Invention**

**[0020]** With reference to Figure 1, there is shown a block diagram of a network data distribution system compatible with the invention. Figure 1 includes a local device 100, a data access network 120, a recommendation server 130 and a web server 150.

**[0021]** The local device 100 may be a client computer that is configured to access the web server 150 and the recommendation server 130 via the data access network 120. The client computer may be, for example, a PC running a Microsoft Windows operating system. The local device 100 includes an output device, such as display 101, and an input device, such as keyboard 102 and / or pointing device 103 (e.g., mouse, track ball, light pen, or data glove). The local device 100 may also be, for example, an Internet appliance, network computer (NC), or a data-enabled device such as a portable digital assistant (PDA), mobile phone, refrigerator, automobile, etc. The particular type of device of the local device 100 is not considered to be important so long as the local device 100 can provide some measure of individual user interactivity with a source of electronic content via a data access network in a client/server fashion.

**[0022]** The data access network 120 provides lower layer network support for the local device 100 to interact with servers in the network data distribution system, including the recommendation server 130 and the web server 150. The data access network 120 may comprise a common or private bi-directional telecommunications network, a public switched telephone network (PSTN), a cable-based telecommunication network, a LAN, a WAN, a wireless network, any of which are coupled with or overlaid by a TCP/IP network (e.g., the Internet or an intranet).

**[0023]** The web server 150 may be of the type known in the art and has the ability to serve web pages to the local device 100, as requested in the manner known in the art. It should be appreciated that the web server 150 is representative of any source of web pages and electronic content available to the local device 100. Thus, for example, the web server 150 could be accessible from the Internet, or it could be a part of an intranet, and represents any number of servers.

**[0024]** The recommendation server 130 is a computer system, such as a server computer. The recommendation server 130 may be considered to represent a number of physical devices which as a group provide the indicated network services. For example, the recommendation server 130 could include a web server plus a database server. The recommendation server 130 transmits certain data to the local device 100 as described further below. The recommendation server 130 may also act as a recipient of certain information transmitted by the local device 100, as described further below.

[0025] Referring now to Figure 2, there is shown a block diagram of another view of the network data distribution system of Figure 1. The system comprises a client 110, the data access network 120, the recommendation server 130 a recommendation database 180 and a page categorization database 140.

[0026] A browser 160 is also shown. A "browser" is software that provides interactive utilization of units of electronic content, such as web pages. The browser 160 may be Microsoft Internet Explorer or Netscape Navigator. The browser 160 may alternatively be a microbrowser used to browse units of WML or HDML based electronic content on a wireless handset. When the local device 100 is connected to the web server 150 through the data access network 120, the user of the local device browses the web server 150 from the local device 100 using the browser 160. The browser 160 need not be stored on the local device 100. The user, from the local device 100, can exercise control over what electronic content is requested and thus output to the output device of the local device 100.

[0027] The client 110 is software operative on the local device 100. The client 110 may be an independent application program, a DLL or other logical grouping of routines. The client 110 need not be stored on the local device 100. The client 110 may be integrated with the browser 160, an operating system, or other software.

[0028] The recommendation database 180 and the page categorization database 140 store and provide data regarding categories, web pages and recommendations. The page categorization database 140 supports category look-up for electronic content. The

recommendation database 180 supports category-based recommendations. Although described herein as separate entities, the recommendation database 180 and the page categorization database 140 may be combined into a single database with appropriate fields and controls, and may be otherwise distributed.

**[0029]** A copy or subset of the recommendation database 180 and the page categorization database 140, referred to as local cache 170, may also be stored in the local device 100 to speed the operation of the client 110. The client 110 and recommendation server 130 may cooperate to update the local cache 170, and to have the recommendation database 180 and the page categorization database 140 accessed when the local cache 170 is inadequate or unavailable. The decision on what, if anything, to place into the local cache 170 depends on such factors as the capabilities of the client 110, the recommendation server 130, the databases 140, 180, and the data access network 120. Decisions on what and how much to store in the local cache 170 may be influenced by factors such as popularity of an object to a particular user and popularity to all or a group of users.

**[0030]** One aspect of the present invention is the use of "categories." A category has two components. First, there is a label associated with the category which in most embodiments is descriptive of the category. Second, there is a scope for the category. The category scopes may be precisely defined, or may be loosely defined. The scopes may be defined through automated and/or manual techniques. Scopes may be defined using principles of linguistics and cognitive science. The particular labels and scopes, and the method of creating the labels

and scopes, is not critical to the invention. Furthermore, the labels and scopes to be used are generally dependant on the embodiment of the invention.

**[0031]** In general, the categories should be logically distinct, though some overlap may be inevitable. The categories should be of human interest, which is itself difficult to precisely define. Just as there are numerous techniques for selecting categories, so too there are numerous techniques for categorizing units of electronic content such as web pages, and for selecting which unit of electronic content to recommend for a given category. In the embodiment currently contemplated, formulation of the page categorization database 140 and the recommendation database 180 involves human input. In the embodiment currently contemplated, the page categorization database 140 comprises domain names and URLs which are selected based upon popularity. Objects (e.g., domain names and URLs which resolve to web pages) in the page categorization database 140 are categorized by parsing the HTML of the corresponding web page, distilling the text of the pages, and deriving a sense of the text of each page. The senses may be made using principles of linguistics and cognitive science. The senses are used to select one or more categories into which the web page fits. The list may include ratings of relevance of a given web page to its categories.

**[0032]** Referring now to Figure 3, there is shown the display 101 having a client window 350 and a browser window 300. The client window 200 is generated and controlled by the client 110. The browser window 300 is generated and controlled by the browser 160. The browser window 300 is familiar to those skilled in the art, so the particulars are not described

further herein. Further information regarding the use of most browsers and their technical specifications is abundantly available.

**[0033]** The browser window 300 includes a display pane 310, an address bar 320 and a title bar 330. The display pane 310 is a region of the browser window 300 wherein the browser 160 causes web pages received by the browser 160 to be displayed. The address bar 320 is another region of the browser window 300. The browser 160 displays URLs in the address bar 320 corresponding to the web page currently displayed in the display pane 310. The user can also enter a URL into the address bar 320, and the browser 160 will attempt to load the web page or other object to which the entered URL points. The address bar 320 may be hidden. However, there is an object associated with the address bar which, in common practice, stores the URL for the currently displayed web page. The primary feature of the title bar 330 is that it displays the title of the browser 160. Another feature of most browsers is that the title bar 330 displays the title of the web page then displayed in the display pane 310.

**[0034]** The client window 350 includes a title bar 351 and a number of operational icons 352, 353 on the title bar 351. The title bar 351 may be used for identifying the client 110. The client window 350 as shown includes a recommendation pane 360. The recommendation pane 360 includes a prompt 361, a category display area 362 and an activation button 363. The client window 350 and the recommendation pane 360 are shown having a conventional rectangular shape. However, the client window 350 and the recommendation pane 360 may define any of a wide variety of regular or irregular shapes.

[0035] The client window 350 is displayed on top of the browser window 300. The client window 350 may be configured to attach to an edge of the browser window 300, and always remain visible and on top of the browser window 300 (persistent). The location of the client window 350 may be predefined, selectable by the user, or selected by a server remotely. In one embodiment, the client window 350 is attached to the title bar 330 of the browser window 300. In other embodiments, the client window 350, or parts of the client window 350, may be integrated into the browser window 300. For example, the title bar 351 of the client window 350 may be eliminated, and the contents of the recommendation pane 360 fixed in the browser window 300.

[0036] The operational icons 352, 353 on the title bar 351 include a close icon 352 and a help icon 353. Activation of the close icon 352 causes client 110 to close the recommendation pane 360, although the title bar 351 of the client window 350 remains displayed. The help icon 353 may be used for providing help to the user. The category display area 362 is used for displaying the label associated with the category of the web page being displayed in the browser display pane 310. The prompt 361 is static text which, when combined with the display in the category display area 362, conveys a message to the user of the availability of a recommendation. The activation button 363 is used by the user to accept the recommendation.

[0037] Variations of the client window 350, and corresponding functionality of the client 110 are within the scope of the invention. The category display area 362 may be a drop down list. In such an embodiment, the drop down list could include all of the categories in which

the current web page falls, and could list sub-categories. The client window 350 may provide a selectable display of the URLs, page names, or site names of the recommended web pages.

### **The Methods of the Invention**

**[0038]** Referring now to Figure 4, there is shown a flow chart of a method of recommending electronic content to a user in accordance with the invention. As will be seen, in contrast with prior art systems, this method may be practiced without any particular information about the user, such as a user profile. Because the method is automatic from a user perspective, a user need not register or provide information before gaining its benefits.

**[0039]** After the method begins (step 405), the client 110 activates on the local device 100. The client 110 may activate automatically, for example when the browser 160 activates (step 410). The process by which the client 110 is installed on the local device 100 is not significant. The client 110 may be provided to users for free or for a fee. The recommendation service of the invention may be provided for free or for a fee. Fees may be assessed through well known payment systems, including through artificial media of exchange such as RocketCash.

**[0040]** Once activated, the client 110 can monitor the browser 160 (step 415). One reason that the client 110 monitors the browser is to know when the user has browsed to a new web page. By monitoring the browser's address bar object, when the browser 160 requests a web page, the client 110 can recognize that the address bar object has changed.

This is only one of many techniques for the client 110 to learn that the current web page has changed.

**[0041]** If there is a new web page in the browser 160, then the category of the newly current web page is determined (step 420). The address bar object stores the URL of the current web page. By copying the contents of the address bar object, the client 110 can use the URL of the current web page as a basis for determining the category of the current web page.

**[0042]** As explained above, the recommendation server 130 has access to the page categorization database 140, which can be copied to the local device 100 and accessed directly by the client 110. Depending on the location and distribution of the page categorization database (stored in the local cache 170 or the full copy 140), the URL from the address bar object may be used to obtain the corresponding category from the categorization database. This may also depend on whether the categorization database 140 stores complete URLs, portions such as domain names, keywords, etc.

**[0043]** Categorization of the current web page may be done "on the fly." For example, the same techniques discussed above for batch categorization to create the page categorization database 140 may be used on an as-needed basis during browsing. A look-up table may be useful in any event to correlate between categories and characteristics of web pages.

**[0044]** Although the category determination can be made by initially copying the URL from the address bar object, other techniques can be used to determine the category of the current web page. For example, techniques such as screen-scraping, data-stream sniffing, and copying other objects used by the browser are available to obtain information about the current web page. This information may be used as described above for categorization.

**[0045]** Once the category of the current web page is known, the client 110 can select a web page to recommend (step 425). There may be only one web page recommended, or a number of web pages may be selected to recommend. The recommendation is automatic (i.e., active), and does not require the user to do anything to obtain the recommendations. Furthermore, recommendations may be made without reference to user profiles of any kind.

**[0046]** To make a recommendation, the client 110 consults the recommendation data in the local cache 170, in the recommendation database 180, or may cooperate with the recommendation server 130 to access the recommendation database 180. In the embodiment currently contemplated, web pages to recommend for each category are selected in advance. This includes some measure of human involvement to refine the recommendations. When needed by the client 110, recommendations may be made in a way that minimizes the chance that a recommendation is made twice. Recommendations may also be made on a preferential basis, and may be made on exchange of consideration (e.g., paid placement). Recommendations can also be made on numerous other factors, including popularity, fit in a category, and relationships.

[0047] Furthermore, recommendations can be made by synthesizing characteristics from a user's historical web browsing. Since the client 110 monitors web browsing, a history of web pages browsed may be maintained and utilized to enhance the recommendations. For example, it may be desirable to not recommend pages which the user has already browsed, which the user has browsed them within a certain period of time, or which are similar to pages the user has browsed or recently browsed.

[0048] It is believed that, for the client 110 to be effective, its use should instill trust in the user. This theory is drawn from the experience of prior art search engines. Thus, although short-term revenues may be enhanced by accepting paid placements which are not particularly relevant to a category, this may ultimately reduce usage of the client 110 because of reduced user trust.

[0049] After the client 110 has obtained one or more recommendations for the current web page (step 425), the client 110 generates a message on the output device 101 which informs the user of the availability of a recommendation (step 430). This may be achieved by displaying the category of the current web page in the category display area 362 (Figure3), and displaying the activation icon 363. The display of the category in the category display area 362 may occur after the category of the current web page has been determined in step 420, or after the recommendation(s) are available in step 425.

[0050] The user then may provide input to the client 110 indicative of the user's desire to activate the recommendation (step 440). If there is more than one recommendation, the

activation icon 363 may be used to select the first recommended web page. The user may be provided with the opportunity to select from a list of recommended web pages, for example with a drop down list. If the user chooses not to accept the recommendation, then browsing continues (step 445), and the client 110 continues to monitor the browser (step 415).

**[0051]** As an alternative to step 440 following step 425, they may be reversed. That is, the user may provide input to the client 110 indicative of the user's desire to activate a recommendation, and then the recommendation may be obtained.

**[0052]** If the user accepts a recommendation, then the client 110 causes the browser 160 to request the recommended web page (step 450). If the user could select from more than one recommendation, then the client 110 causes the browser 160 to request the recommended web page which the user selected. The web browser then requests and loads the recommended web page. The recommended web page may be displayed in the same browser window 300 as the current web page, or may be displayed in a new or other window.

**[0053]** After the user has selected a first recommended web page (step 440), the user may continue to accept recommendations from the same category (step 455). This step 455 may be performed in a number of ways. As mentioned, in step 425 several web pages may be selected for recommendation. In such a case, the client 110 maintains a list of recommendations and the user may select a next recommendation by activating the activation icon 363. The activation icon 363 may change appearance to reflect that more recommendations are available. As an alternative to selecting multiple web pages to

recommend in step 425, after a recommended web page has been loaded, it can be treated as the current web page, with control returning to step 415. By providing successive recommendations, the user may be provided with an electronic tour of web pages which are likely to be of interest.

**[0054]** In the currently contemplated embodiment, the client 110 continues operating so long as the browser 160 is active. Thus, the client window 350 is never completely closed. Since the client window 350 may be embodied in other forms, or eliminated as described above, manners of hiding operative features of the client 110 or closing the client 110 are within the scope of the invention.

**[0055]** The client 110 may provide additional avenues for making recommendations of web pages to the user. The user may be prompted to enter or select one or more key words, categories or linguistic senses, provided as a whole or limited by relevant criteria as discussed above. Recommendations may then be made from the user input alone or in conjunction with analysis of the current web page.

**[0056]** Although exemplary embodiments of the present invention have been shown and described, it will be apparent to those having ordinary skill in the art that a number of changes, modifications, or alterations to the invention as described herein may be made, none of which depart from the spirit of the present invention. All such changes, modifications and alterations should therefore be seen as within the scope of the present invention.